INTRODUCTION

The Ocean Drilling Program (ODP) was an international partnership of scientists and governments who joined together to explore the structure and history of the earth beneath the ocean basins. The central purpose of ODP was to provide core samples, downhole measurements and other scientific measurements to characterize the oceans' basins. The data generated are leading to a better understanding of the processes of plate tectonics, earth's crustal structure and composition, conditions in ancient oceans, and changes in climate through time.

As the Science Operator, Texas A&M University (ODP/TAMU) had the responsibility to collect cores from the oceans basins, provide adequate facilities for analyses of the cores, and assure the preservation of the core and scientific data collected by shipboard scientists. This program was very successful – ODP repositories hold over 220 km of core and the databases contain millions of analytical measurements made on core sections and samples. By providing the drilling platform, consistent drilling procedures, on-board laboratories equipped with analytical equipment, and standard data collection procedures, the ODP has created a suite of analyses from locations all over the world's oceans that are directly comparable.

PRIME SCIENTIFIC DATA

One of the primary tasks of the science operator was to collect, archive and disseminate scientific data collected on the core during the cruises. The table below contains a list of the prime data types – analyses performed on the ship during a cruise with standard data collection procedures on all core when practical.

Most of the scientific measurements are available using web-based queries of the Janus database, an Oracle database created to serve data to the scientific community as well as maintain the prime data in a permanent archive. The Janus database is now maintained by IODP/TAMU – the United States Implementing Organization of the Integrated Ocean Drilling Program. The Database website maintains a set of queries that allows the user to extract data.

This document provides a general description of the prime scientific data types. For each data type the following information will be provided:

1) Description of data collection procedures and the equipment used;
2) How the data were archived before the Janus database existed;
3) How the data are stored in the Janus database;
4) Discussion of data quality. Factors that may affect the quality of the data.

More detailed information about ODP data collection can be obtained from the technical notes which describe data collection for each of the different laboratories on the JOIDES Resolution. In addition, there are the Explanatory Notes at the beginning of each Initial Report volume, Laboratory Officer's report for each leg, laboratory technicians' reports, and logsheets that can provide additional information. This information can be requested from the IODP/TAMU Data Librarian.

Prime Scientific Data	**Data Availability**	**Leg first collected**
Operations and Core Summary		
Leg summary and details	Janus	101
Site summary	Janus	101
Hole operations summary	Janus	101
Core summary	Janus	101
Lithology and Stratigraphy		
Visual Core Description	Janus – primedata images Initial Report volumes	101
Core Photos	Janus – scanned photographs	101
Reentry Video	IODP Data Librarian	
Digital Scanning of Sections	Janus	198
Spectral reflectance and colorimetry	Janus	154
Paleomagnetism	Janus	103
Tensor	Janus	174
Smear Slides and Thin Sections	IODP Data Librarian Initial Reports volumes	101
Biology and Stratigraphy		
Paleontology	Janus *	101
Age Profile	Janus *	101
Microbiology	IODP Data Librarian	185
Geochemistry and Mineralogy		
Carbon / Carbonate Content	Janus	101
Gas Chromatography	Janus	101
ICP - AES	Janus	187
Interstitial Water	Janus	101
Rock Evaluation	Janus	101
X-ray Fluorescence	Janus	106
X-ray Diffraction	Janus	106
Physical Properties		
Gamma Ray Attenuation	Janus	101
Index Properties	Janus	101
Magnetic Susceptibility	Janus	101
Natural Gamma Radiation	Janus	150
P-wave Velocity (logger)	Janus	108
Shear Strength	Janus	101
Sonic Velocity (samples)	Janus	101
Thermal Conductivity	Janus	101
Temperature and Geophysics		
Downhole Temperature	Janus	104
Underway Geophysics and Bathymetry	National Geophysical Data Center IODP Data Librarian	101
Seismic Surveys	IODP Data Librarian	101

* Migration of these data to be completed by 2007. Contact IODP Data Librarian for information.

General Definitions

LEG

The Ocean Drilling Program started numbering the scientific cruises of the *JOIDES Resolution* at Leg 101. (Leg 100 was a trial run of the modified drilling ship.) A leg was nominally two months duration. The shipboard science party typically consisted of 25 scientists drawn from universities, government and industry around the world. During the 18+ years of the ODP, there were 110 cruises on the *JOIDES Resolution,* more commonly known as the *JR*.

SITE

A site is the location where one or more holes were drilled while the ship was positioned over a single acoustic beacon. The *JR* visited 656 unique sites during the course of the ODP. Some sites were visited multiple times, including some sites originally visited during the Deep Sea Drilling Program for a total of 673 site visits.

HOLE

Several holes could be drilled at a single site by pulling the drill pipe above the seafloor, moving the ship some distance away and drilling another hole. The first hole was designated ‘A’ and additional holes proceeded alphabetically at a given site. Location information for the cruise was determined by hole latitude and longitude. During ODP, there were 1818 holes drilled or deepened.

CORE

Cores are numbered serially from the top of the hole downward. Cored intervals are up to 9.7 m long, the maximum length of the core barrel. Recovered material was placed at the top of the cored interval, even when recovery was less than 100%. More than 220 km of core were recovered by the ODP.

CORE TYPE

All cores are tagged by a letter code that identifies the coring method used. Some of the more common core types are listed below.

- H – APC - The Advanced Piston Core is a hydraulically-actuated piston corer designed to recover undisturbed core from soft sediments. It is designed to be delivered through the drill string to the sediment to be cored whether at the sea floor or hundreds of meters below it.
- X – XCB - The Extended Core Barrel is designed to recover core samples from soft to medium formations. Typically, the XCB is deployed upon APC refusal, i.e., when a formation becomes too stiff to piston core. The XCB relies on rotation of the drill string to advance the hole and cut the formation.
- R – RCB - The Rotary Core Barrel is designed to recover core samples from medium to hard formations. The RCB relies on rotation of the drill string to advance the hole and cut the core. The core bit trims the sample.
- W – WASH - When washing down a hole, drilling rates are higher if a core barrel is used rather than blocking the drill bit with a center bit device, even though no core is desired. The driller may wash down a hole as many meters as desired without retrieving a core.
- G - A ghost core is one whose contents come from an already drilled part of the hole, i.e., the extent of a ghost core lies completely within the drilled or cored portion of a hole.
- P – PCS - The Pressure Core Sampler is capable of retrieving core samples from the ocean floor while maintaining near in-situ pressures up to 689.7 atmospheres (10,000 psi).

- M – MISC - This represents material that could not be labeled with a standard core type. This category includes limited numbers of cores that are recovered using experimental drilling methods which, once they are established, are assigned their own core type.
- Z – DIAM - Diamond Coring bits are being developed to enhance recovery of core in hard rock.
- N – MDCB - The Motor Driven Core Barrel is a wireline-retrievable coring system designed for a two-fold purpose. It allows a single bit APC/XCB holes to be extended to greater depths and into more indurated formations. The MDCB can also improve recovery in difficult formations.
- 0 – 9 -- Interval was drilled, not cored.

SECTION

Cores are cut into 1.5 m sections in order to make them easier to handle. Sections are numbered serially, with Section 1 at the top of the core. Most of the scientific measurements were made on sections or discrete samples taken from the sections. Samples and measurement intervals are given in centimeters from the top of each section. After being cut into sections, several whole-core measurements were made, and then the core was split into working and archive halves. The archive halves were used for the visual descriptions, paleomagnetics and photography. The working halves were sampled for shipboard and shore-based studies.

Core Handling, Curation, Sampling and Analysis

As soon as core was brought on deck, a paleontology sample was usually taken from the core catcher in order to get an initial age assessment. It was then put into a long rack where vacutainer gas samples could be taken from voids. Sections were marked, labeled and then cut. Whole-round samples were taken. Each section was sealed with color-coded caps glued on the top and bottom – a blue cap for top of section, a clear cap for the bottom, and a red cap for where a whole-round sample was taken.

The sections were moved to the core lab where they were labeled with an engraver to mark the full identification of the section. The length of core was measured for each section and core catcher; the lengths were entered into the database. The cores were equilibrated to room temperature, run through the MultiSensor Track (MST) which included Gamma Ray Attenuation, Magnetic Susceptibility, Natural Gamma Radiation, and P-wave Velocity, and thermal conductivity measurements were taken.

Now the sections were split into working and archive halves. The softer sections could be split with a wire, the more indurated sections were cut with a saw. Harder rock cores were split with a band saw or diamond saw. The archive halves were wrapped and taken to the core description area where scientists completed a visual core description. The archive halves were analyzed using the pass-through cryogenic magnetometer. The working halves were available for sampling. Routine samples for the shipboard laboratory analyses were taken. In addition, scientists who had requested personal samples could take samples for their own research. These could be analyzed in the shipboard laboratories, or taken back to the scientists' laboratories. All samples were entered into the database and included location information (leg, site, hole, core, core type, section, top and bottom intervals) and volume of sample.

Data Acquisition

Most of the analytical laboratories on the *JR* were equipped with analytical equipment that interfaced to computers. Early in the program, these programs were fairly basic, but as technology advanced, the data acquisition programs became more advanced. Sections and samples were identified by barcode after Leg 171, but the implementation of barcode readers on each of the analytical systems had not yet been accomplished.

Sources of errors. Most data collection in the *JR* laboratories involved at least one manual data entry event - sample entry into database; section or sample information entry to data acquisition program; data entry into databases; or reformatting of data tables for the Initial Report. Core and section information were checked very closely, but verification of every sample, every analysis, and every run can be an enormous problem during the operational time of a cruise. Operator entry or typographical errors probably account for the largest number of errors.

Data Archive

Data Archive at ODP before Janus.

Several of the prime data types have been captured digitally since the beginning of the program. Files were created on computers in the laboratories and transferred to the central computer for compilation and transfer back to shore at the end of the leg. ODP/TAMU retained many of the original 'raw' files on an active server, each new leg added after data returned from the ship. These files were maintained as an archive and were made available to the scientific community on request to the ODP Data Librarian.

During the early part of ODP, the science operator recognized the need for a data management system to help storage and retrieval of the scientific data. The system that was chosen was System 1032 by Computer Corporation of America or S1032. Most of the scientific data collected on samples, e.g., index properties, carbonate analyses, etc., were loaded into S1032 and the original files not saved on line. The media that were used to transfer data between ship and shore are obsolete, so retrieval of the original raw files for some datasets is no longer possible.

Throughout the program, database technologies were evolving, and ODP scientists and staff were continually looking for more efficient and effect ways to capture the information collected on the ship. S1032 was used for a variable length of time depending on the data set. Some effort was made to move some data sets, e.g., index properties and paleontology, into 4D, a relational database running on the Macintosh computer system. For some data types, it was decided to archive the original files. There were problems with these database engines including software upgrades that could not read databases created with earlier versions, and some versions of these databases were easily corrupted.

Data Archive at ODP after Janus

The continuing improvement in database technology and the explosion of scientific databases accessible over the Internet convinced ODP/TAMU to create a database that would provide a permanent archive for the scientific data and could be accessed by the public through an Internet interface. Janus, an Oracle-based relational database, was created and has become the repository for all core information, sample information, and most of the shipboard scientific data. Janus became operational on Leg 171, January 1997. From that point, the scientific data generated on board the *JR* were uploaded to Janus; but the original data files were still brought back to shore and archived. Janus Web was created to provide standard data queries that allowed the public to look at or download data from Janus.

Janus Data Management and Verification

Before Janus became operational, verification of the scientific data collected during a cruise was the responsibility of the leg scientific party. Minimal checking or verification was done after the data were brought back to shore. However, when Janus became operational on Leg 171, more complete verification of leg-associated scientific data became necessary. The relational characteristics of the Janus database require more active oversight by database managers to ensure that integrity of the data is maintained. ODP data verification procedures include:

- Data Acquisition - Confirmation that the data acquisition software is recording all the data necessary to document each analysis, and ensure that Janus has a place to put every piece of information that should be saved.

- Upload - Testing the upload software to verify that each datum is being put into the correct field.

- Leg Data - Checking data collected on a leg to help ensure that data are linked correctly to the sections and samples on which the analyses were performed.

- Web Queries - Confirmation that web queries return the correct information to the user. If the relationships in the database are not correct, critical data are missing, or the query is not written properly, data values can incorrectly be associated with other samples or sections – not the core material on which the analyses were done, or the query may not work at all.

Data checking was by far the most difficult and sensitive task for database personnel. The following procedures were instituted to verify that the data from a leg was properly entered into the Janus database, and at the same time, reassure the scientific community that database personnel were not changing data without the knowledge and permission of the scientific party.

1. Use logsheets and lab notes whenever available. Logsheets had been used in many of the laboratories since Leg 100, and most logsheets were returned to

shore for archival early in the program. Most of the data entry into S1032 and data verification was done on the ship by the scientific party, so the practice of sending the logsheets to shore at the end of the leg was discontinued. After Janus became operational and the need for database oversight was recognized, the logsheets were returned to shore to aid the verification procedures.

2. Check the Laboratory Technician reports, Explanatory Notes and other communication to document any problem that could affect the quality of the data.
3. Interact with the scientific party when a question about any aspect of the data collection was identified. Database personnel attended the Leg Post-Cruise meeting in which the final editing was done on the Initial Report volume for that leg. Database personnel could interact with the scientists and resolve problems either in the database or the Initial Report data tables.

Unfortunately, the verification of the pre-Janus migrated data was a more difficult problem. Logsheets and lab notebooks were not available for most of the legs' data sets. The important link of interacting with the scientists and technicians who collected the data was mostly missing. The main resource was the Initial Report volumes. When there was a discrepancy, all available information was studied in order to determine the nature of the discrepancy. Data in the database were only changed when there was compelling evidence that data in the Initial Report table were wrong. Some of the data-specific verification techniques and common sources of errors are discussed in the data reports.

References

A comprehensive list of all the reports, logsheets, notebooks, conversations, hand-written notes, etc., used to compile the information for the data summary reports is impossible. For additional information about any aspect of the ODP data collection procedures, please contact the IODP Data Librarian.

IODP Data Librarian

Integrated Ocean Drilling Program
Texas A&M University
1000 Discovery Drive
College Station, Texas 77845

database@iodp.tamu.edu